A Deep Learning Framework for Fish Disease Classification Using EfficientNetB7 and Gaussian-Inhibited Spatial Attention

Hoang-Tu Vo[1,2], Huu-Hoa Nguyen[1], Vinh Dinh Nguyen[2]
[1] College of Information and Communication Technology, Can Tho University, Vietnam
[2] Information Technology Department, FPT University, Cantho city, Vietnam
Correspondence: Huu-Hoa Nguyen, nhhoa@ctu.edu.vn
Communication: received 30 June 2025, revised 28 August 2025, accepted 01 October 2025
Digital Object Identifier: https://doi.org/10.32913/mic-ict-research.v2025.n3.1400

Abstract: Accurate classification of fish diseases is essential for maintaining the health and productivity of aquaculture systems. This study proposes a novel deep learning framework that combines the EfficientNetB7 architecture with a custom attention mechanism called Neural Spatial Attention with Gaussian Inhibition (NSAG). The NSAG module integrates both channel and spatial attention pathways, enhanced by a Gaussian inhibition mechanism that sharpens focus on the most relevant spatial regions. To validate the effectiveness of the proposed approach, we performed extensive experiments on a dataset of fish disease images. Comparative evaluations were performed across five configurations: (1) fine-tuned EfficientNetB7, (2) EfficientNetB7 combined with Squeeze-and-Excitation (SE) attention, (3) EfficientNetB7 with spatial attention, (4) EfficientNetB7 integrated with Convolutional Block Attention Module (CBAM), and (5) our proposed EfficientNetB7 + NSAG architecture. All models were trained under identical conditions for fair comparison. The experimental results provide evidence that the proposed model consistently outperforms the competing baselines, achieving an accuracy exceeding 97.94%, with an improvement of nearly 3% over the CBAM-based model. These findings demonstrate the potential of NSAG in boosting attention learning and improving the classification of fish diseases.

Keywords: *Fish disease classification, EfficientNetB7, attention mechanism, spatial attention, CBAM, SE attention, Gaussian inhibition.*

I. Introduction

The fast-growing development of aquaculture has led to increased concerns regarding the early recognition and diagnosis of fish diseases, which can significantly impact productivity, sustainability, and food security. Traditional disease identification methods are labor-intensive, time-consuming, and often require expert intervention. With the growing availability of image-based fish disease datasets and advances in artificial intelligence, deep learning (DL) has proven as a promising method for automating disease classification tasks with high accuracy. Among deep learning architectures, EfficientNet has demonstrated remarkable performance in various computer vision tasks due to its compound scaling and parameter efficiency. In particular, EfficientNetB7 offers a strong balance between performance and computational efficiency, making it a suitable backbone for medical, agricultural, and aquacultural image classification. However, despite its strength in feature extraction, standard EfficientNet lacks an explicit attention mechanism to focus on the most informative spatial and channel-wise regions, which is particularly important in fine-grained classification problems such as differentiating among visually similar fish diseases. To address this limitation, attention mechanisms such as SE, Spatial Attention, and CBAM have been proposed and widely adopted. While these modules improve model focus, they often fail to suppress irrelevant background noise in a biologically plausible manner.

In this research, we introduce a new attention module called Neural Spatial Attention with Gaussian Inhibition (NSAG), which enhances both channel and spatial attention while selectively suppressing less informative regions using a Gaussian-centered inhibition strategy. We integrate the proposed NSAG module into the EfficientNetB7 architecture and evaluate its effectiveness on a fish disease image dataset. A series of comparative experiments are conducted, including fine-tuning baseline EfficientNetB7, combining it with SE attention, spatial attention, and CBAM. The results show that our proposed NSAG-based model achieves superior classification performance, exceeding 97.94% accuracy and outperforming all compared baselines, with an improvement of nearly 3% over the CBAM-based model.

The remainder of this paper is organized as follows: Section 2 shows a review of related work on fish disease classification using DL techniques, with a focus on existing approaches and attention mechanisms. Section 3 introduces the proposed model architecture, including the NSAG Attention Module and the overall model formulation. Section 4 describes the experimental configuration and discusses the obtained results, covering evaluation metrics and comparative performance. Finally, Section 5 concludes the study and proposes future research directions.

II. Related works

In recent years, DL has shown significant promise in addressing challenges in fish disease diagnosis through automated image classification [1–5]. Convolutional neural networks (CNNs) have become the backbone of such applications as they enable end-to-end learning of multi-level features directly from raw pixel inputs, bypassing manual feature extraction. Early approaches typically employed conventional CNNs such as VGGNet [6] or ResNet [7], achieving reasonable accuracy across various tasks [8–13] but often suffering from overfitting or excessive model complexity when applied to small or imbalanced datasets.

To address model efficiency and scalability, the EfficientNet [14] family was introduced and has gained popularity in medical [15–17] and agricultural [18, 19] imaging tasks. EfficientNet leverages compound scaling to optimize the architecture across depth, width, and resolution dimensions, providing a more parameter-efficient architecture. Several studies have employed EfficientNet variants for disease classification in crops [20–22], livestock [23, 24], and aquatic species [25], reporting improved accuracy over traditional CNNs. Alongside architectural advancements, attention mechanisms have been widely adopted to enhance model focus and interpretability. The authors in [26] introduced channel attention by adaptively recalibrating feature responses. The CBAM [27] further improved upon this by combining both channel and spatial attention in a sequential fashion. These attention mechanisms have demonstrated significant performance gains in fine-grained image classification tasks, including disease detection in fish and plants. Recent work, for example, has integrated CBAM into EfficientNetB6 for fish disease recognition, achieving an accuracy of 99.45% on specialized aquaculture datasets [28]. Other studies have applied CBAM or SE attention in medical image analysis, particularly histopathology and microscopic diagnostics, where attention to fine details is critical [29]. However, traditional attention mechanisms tend to distribute focus across multiple regions, which may result in the inclusion of irrelevant background noise. While some biologically inspired models have introduced inhibition concepts, few have employed Gaussian-centered inhibition to sharpen spatial focus in a learnable attention context.

In this study, we propose a novel attention module-Neural Spatial Attention with Gaussian Inhibition (NSAG)-which combines standard channel and spatial attention mechanisms with a biologically motivated inhibition process to refine spatial focus. This study explores the integration of a Gaussian-centered spatial inhibition mechanism within an attention framework for fish disease classification-an approach that has received limited attention in previous literature. Comparative experiments with SE, spatial attention, and CBAM demonstrate that the proposed NSAG module leads to superior classification performance, achieving 97.94% accuracy.

III. Research Methodology

This section describes the proposed deep learning model for fish disease classification, which integrates the EfficientNetB7 backbone with a custom attention module named Neural Spatial Attention with Gaussian Inhibition (NSAG). The design is motivated by the need to enhance feature discrimination in fine-grained classification tasks, where disease patterns may be hard to distinguish and spatially localized.

1. Overall Architecture

The proposed model (Fig.1) leverages EfficientNetB7 as the base feature extractor due to its superior parameter efficiency and accuracy. To capture multi-scale discriminative features, we extract intermediate feature maps from three key blocks in the EfficientNetB7 architecture: block2b_add (low-level features), block4b_add (mid-level features), and block6b_add (high-level features). Additionally, we obtain the final output of EfficientNetB7 as a global feature representation.

Each of these four feature maps is passed through the NSAG attention module to refine spatial and channel-wise responses. To unify feature dimensions, a 1×1 convolution is applied to reduce the channel size to 64. Spatial resolutions are standardized using UpSampling2D and Resizing layers to a common size of 56×56. All processed feature maps are then fused using element-wise addition. Finally, the fused feature is passed through a global average pooling layer, followed by a dropout layer and a fully connected softmax classifier to predict the disease class.

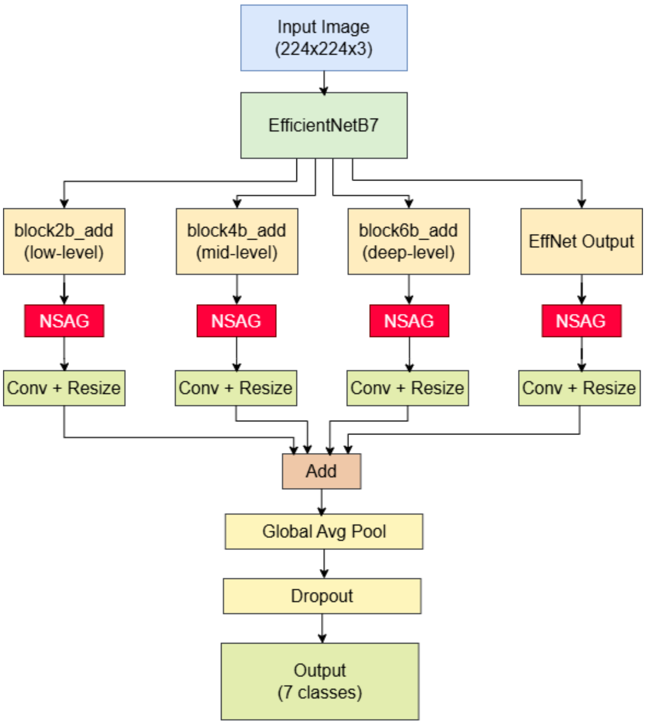


Figure 1. The overall methodology for the classification of fish disease

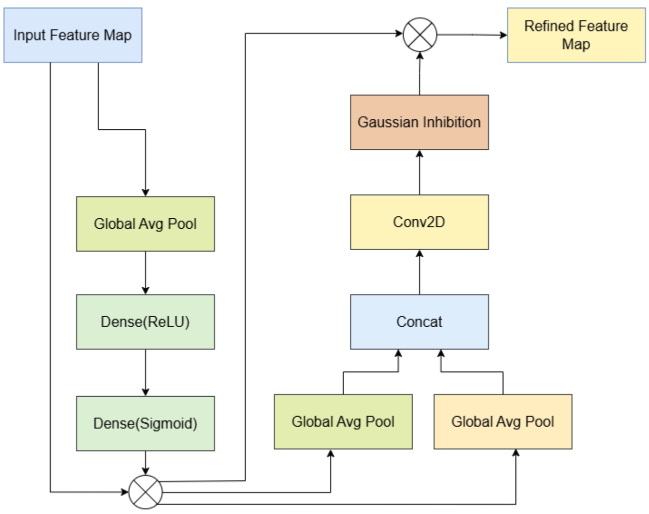


Figure 2. Neural Spatial Attention with Gaussian Inhibition (NSAG)

2. NSAG Attention Module

The proposed Neural Spatial Attention with Gaussian Inhibition (NSAG) module consists of two sequential components: channel attention and spatial attention with Gaussian inhibition (Fig.2).

a) Channel Attention

With a provided input feature map $F \in \mathbb{R}^{H\times W\times C}$, we first apply AvgPool to obtain a channel descriptor $z \in \mathbb{R}^{C}$. This descriptor is fed into a two-layer fully connected network with a reduction ratio r, followed by a sigmoid activation function:

$$z' = \sigma(W_2 \cdot \mathrm{ReLU}(W_1 \cdot z)) \tag{1}$$

where $W_1 \in \mathbb{R}^{\frac{C}{r}\times C}$ and $W_2 \in \mathbb{R}^{C\times\frac{C}{r}}$. The resulting attention vector z' is reshaped to $(1, 1, C)$ and multiplied element-wise with the original input tensor F to obtain the channel-refined feature map F'.

b) Spatial Attention with Gaussian Inhibition

To capture spatial attention, we apply both AvgPool and MaxPool operations along the channel axis on F', concatenate the two resulting 2D maps, and pass them through a 7×7 convolution layer with sigmoid activation:

$$M_s = \sigma\left(\mathrm{Conv}_{7\times7}\left([\mathrm{AvgPool}(F'); \mathrm{MaxPool}(F')]\right)\right) \tag{2}$$

To further refine the attention map and suppress background noise, we introduce a Gaussian-centered inhibition mechanism. Specifically, we compute the spatial maximum response and apply a soft inhibition function as follows:

$$M'_s = M_s - \alpha(M_s - \max(M_s)) \tag{3}$$

where $\alpha \in [0, 1]$ is a tunable inhibition strength factor. The final output of the NSAG module is obtained by multiplying the refined feature map F' with the inhibited spatial attention map M'_s:

$$F'' = F' \otimes M'_s \tag{4}$$

where $\otimes$ denotes element-wise multiplication.

3. Overall Model Formulation

Let F_i denote the intermediate feature maps extracted from selected blocks of EfficientNetB7, where $i \in \{2b, 4b, 6b, \mathrm{final}\}$. Each feature map F_i is refined using the proposed NSAG module:

$$F'_i = \mathrm{NSAG}(F_i) \tag{5}$$

To unify spatial dimensions across all feature levels, each refined feature map is resized to a fixed resolution using upsampling and interpolation:

$$\hat{F}_i = \mathrm{Resize}(F'_i) \tag{6}$$

The fused feature representation is computed via element-wise summation:

$$F_{fused} = \sum_{i=1}^{4} \hat{F}_i \tag{7}$$

Next, we apply global average pooling (GAP), followed by dropout and a fully connected softmax layer for final classification:

$$\hat{y} = \text{Softmax}(\text{FC}(\text{Dropout}(\text{GAP}(F_{\text{fused}})))) \quad (8)$$

Here, $\hat{y} \in \mathbb{R}^C$ represents the predicted probability distribution over the C fish disease classes.

IV. Experiment Results

In this section, we show the experimental results comparing the proposed EfficientNetB7+NSAG model against several baseline and attention-enhanced models for fish disease classification. The evaluation was conducted using a curated dataset of fish images annotated by disease type. All models were trained and evaluated under identical conditions for fair comparison.

1. Experimental Setup

We conducted experiments using the *Freshwater Fish Disease Aquaculture in South Asia* dataset [30], which was developed to support deep learning-based disease identification in aquaculture. The dataset provides a total of 2,451 high-resolution skin images categorized into seven classes, including Aeromoniasis, Bacterial Gill Disease, Bacterial Red Disease, Fungal Saprolegniasis, Parasitic Diseases, Viral White Tail Disease, and Healthy Fish. The data was randomly divided into 80% training, 10% validation, and 10% testing subsets. All images were resized to 224×224 pixels and normalized to the $[0, 1]$ range. During training, standard data augmentation techniques such as horizontal flipping, random rotation, and zooming were applied to enhance model generalization and reduce overfitting. Sample images from the dataset are illustrated in Figure **??**.

Figure 3. Examples of images contained in the dataset

We fine-tuned the EfficientNetB7 backbone pretrained on ImageNet, using categorical cross-entropy loss and the Adam optimizer.

2. Comparison Models

To evaluate the effectiveness of the proposed NSAG module, we compared the following models:

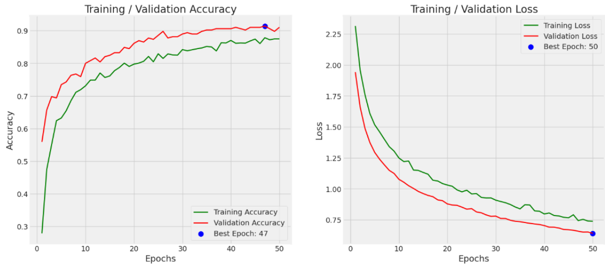
- **Baseline EfficientNetB7**: Pretrained EfficientNetB7 without any attention module.
- **EfficientNetB7 + SE**: Incorporating Squeeze-and-Excitation (SE) blocks after key feature layers.
- **EfficientNetB7 + Spatial Attention**: Applying only spatial attention without channel recalibration.
- **EfficientNetB7 + CBAM**: Combining both channel and spatial attention sequentially using CBAM.
- **EfficientNetB7 + NSAG**: The proposed model with channel-spatial attention and Gaussian inhibition.

Figure 4 presents the training and validation accuracy/loss curves for various model configurations: (a) Baseline EfficientNetB7, (b) EfficientNetB7 integrated with SE attention, (c) EfficientNetB7 with spatial attention only, (d) EfficientNetB7 enhanced by CBAM, and (e) the proposed EfficientNetB7 + NSAG model. Among all variants, the proposed model demonstrates faster convergence and achieves superior validation accuracy with lower validation loss, highlighting its effectiveness over existing attention mechanisms.

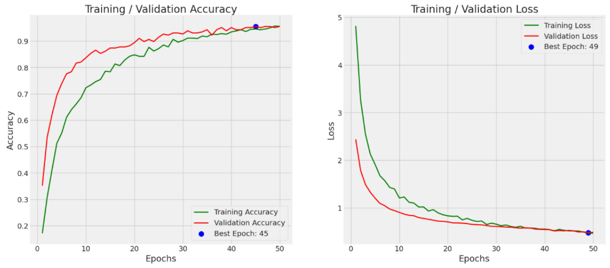
As shown in Table I, the baseline EfficientNetB7 model achieved a strong accuracy of 86.01%. Applying only spatial attention led to slightly better results (90.53%), suggesting that spatially localized features are critical for identifying disease regions.

When enhanced with SE attention, the performance improved to 93.00%, indicating that channel-wise recalibration contributes positively to fish disease classification. The CBAM-integrated model, which combines both channel and spatial attention, further boosted the performance to 95.06%. The proposed EfficientNetB7 + NSAG model outperformed all other configurations, achieving the highest accuracy of **97.94%**, with an improvement of nearly 3% over the CBAM-based model. This demonstrates the effectiveness of the NSAG module in refining both channel and spatial features while suppressing irrelevant regions through Gaussian inhibition. Additionally, consistent improvements were observed across all evaluation metrics (precision, recall, and F1-score), confirming the robustness of the proposed attention mechanism in fish disease classification. Table 2 provides a detailed breakdown of the precision, recall, and F1-score for each fish disease class as predicted by the proposed EfficientNetB7 + NSAG model.

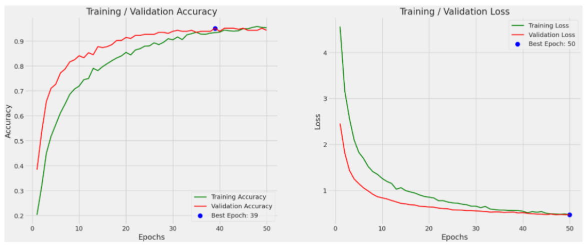
Fig. 5 illustrates the confusion matrices obtained from different network configurations, including the baseline EfficientNetB7 and its variants integrated with SE, Spatial, CBAM, and the proposed NSAG attention modules. As



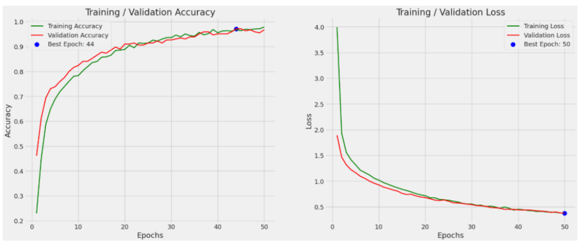
(a) Baseline EfficientNetB7



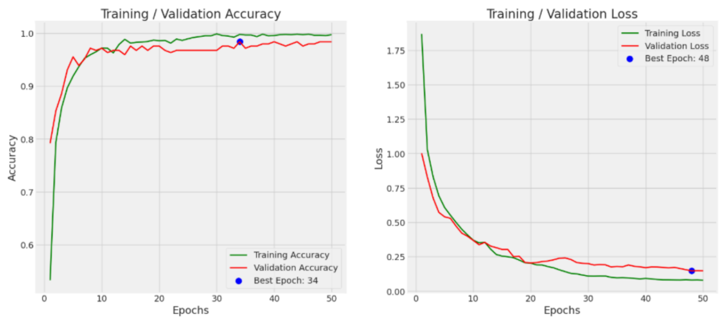
(b) EfficientNetB7 + SE



(c) EfficientNetB7 + Spatial



(d) EfficientNetB7 + CBAM



(e) EfficientNetB7 + NSAG

Figure 4. Comparison of Training and Validation Accuracy/Loss for EfficientNetB7 variants: (a) Baseline; (b) +SE; (c) +Spatial Attention; (d) +CBAM; (e) +NSAG.

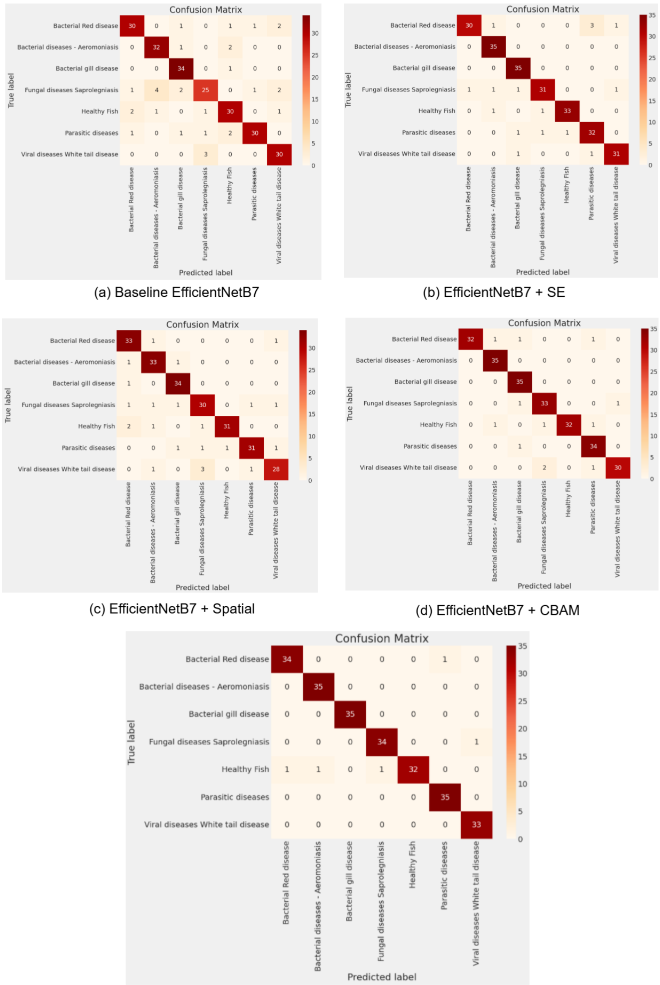
TABLE I
PERFORMANCE COMPARISON OF DIFFERENT ATTENTION MECHANISMS

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Baseline EfficientNetB7	86.01	86.19	85.97	85.88
EfficientNetB7 + SE	93.00	93.31	92.96	93.01
EfficientNetB7 + Spatial	90.53	90.68	90.49	90.49
EfficientNetB7 + CBAM	95.06	95.29	95.03	95.04
EfficientNetB7 + NSAG	**97.94**	**97.97**	**97.96**	**97.93**

TABLE II
PRECISION, RECALL, AND F1-SCORE FOR DIFFERENT FISH DISEASE CLASSES USING THE PROPOSED EFFICIENTNETB7 + NSAG MODEL

Class	Precision	Recall	F1-Score	Support
Bacterial_Red_Disease	0.9714	0.9714	0.9714	35
Aeromoniasis	0.9722	1.0000	0.9859	35
Bacterial_Gill_Disease	1.0000	1.0000	1.0000	35
Fungal_Saprolegniasis	0.9714	0.9714	0.9714	35
Healthy_Fish	1.0000	0.9143	0.9552	35
Parasitic_Diseases	0.9722	1.0000	0.9859	35
Viral_White_Tail_Disease	0.9706	1.0000	0.9851	33
Accuracy	-	-	0.9794	243
Macro Avg	0.9797	0.9796	0.9793	243
Weighted Avg	0.9798	0.9794	0.9792	243

shown in Fig. 5(a), the baseline EfficientNetB7 exhibits misclassifications across multiple disease classes, particularly between Healthy Fish and Viral White Tail Disease, leading to relatively lower performance. Incorporating SE (Fig. 5(b)) and Spatial attention (Fig. 5(c)) improves the class-wise recognition, yet some confusion remains, especially in distinguishing Healthy Fish from Fungal Saprolegniasis. The CBAM-enhanced model (Fig. 5(d)) further reduces these misclassifications, yielding more consistent predictions across disease categories. Finally, the proposed EfficientNetB7 + NSAG model (Fig. 5(e)) demonstrates superior discrimination ability with almost perfect classification, as evidenced by the strong diagonal dominance and minimal off-diagonal errors. These results confirm that the NSAG attention mechanism significantly improves feature representation, thereby enhancing robustness and classification accuracy for fish disease recognition.



(a) Baseline EfficientNetB7

(b) EfficientNetB7 + SE

(c) EfficientNetB7 + Spatial

(d) EfficientNetB7 + CBAM

(e) EfficientNetB7 + NSAG

Figure 5. Comparison of Confusion Matrices for EfficientNetB7 variants: (a) Baseline; (b) +SE; (c) +Spatial Attention; (d) +CBAM; (e) +NSAG.

3. Model Interpretability via Grad-CAM

To further understand how the proposed model makes classification decisions, we employed Grad-CAM to visualize the spatial attention of the network. Fig. 6 and Fig. 7 illustrate the Grad-CAM visualizations generated by the proposed EfficientNetB7 + NSAG model for samples of Bacterial Red Disease and Bacterial gill disease. In both cases, the highlighted red regions correspond to areas of high importance identified by the model during the classification process.

The heatmaps indicate that the model effectively focuses on disease-relevant regions such as skin lesions, fin abnormalities, and localized discolorations, which are critical cues in diagnosing fish diseases. These visualizations support the interpretability of the model and confirm that the NSAG attention module guides the network toward semantically meaningful regions, enhancing both prediction accuracy and trustworthiness.

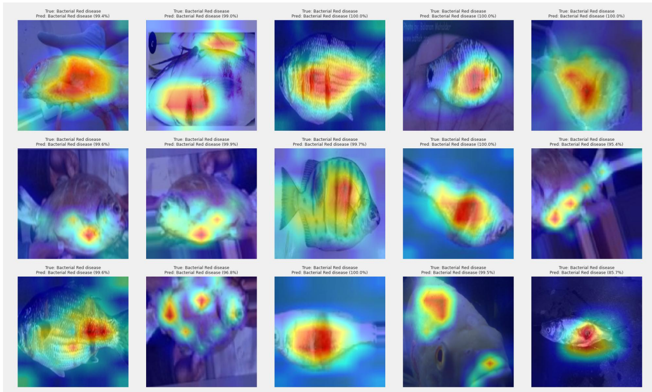


Figure 6. Grad-CAM visualization from the EfficientNetB7 + NSAG model on a sample of **Bacterial Red Disease**. Red regions indicate areas of high importance used by the model for classification.

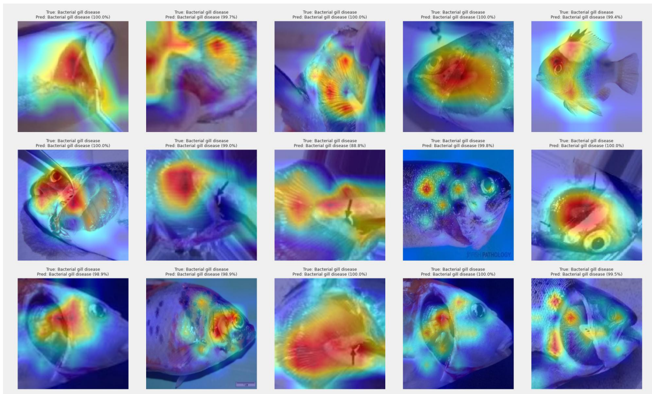


Figure 7. Grad-CAM visualization from the EfficientNetB7 + NSAG model on a sample of **Bacterial gill disease**. Red regions indicate areas of high importance used by the model for classification.

V. Conclusion

In this study, we proposed an enhanced DL architecture for fish disease classification by integrating the NSAG (Neural Spatial Attention with Gaussian Inhibition) module into the EfficientNetB7 backbone. The NSAG module combines both channel and spatial attention, while incorporating a Gaussian inhibition mechanism to suppress irrelevant background regions and enhance feature localization. Experimental evaluations on a curated dataset of fish disease images demonstrated that the proposed model outperforms several baseline and attention-based architectures, including SE, spatial-only, and CBAM modules. The proposed EfficientNetB7+NSAG obtained an accuracy of 97.94%, along with superior precision, recall, and F1-score. These results highlight the effectiveness and robustness of the NSAG module in improving disease recognition under challenging visual conditions. A limitation of the current NSAG module is its use of a fixed 7×7 convolution kernel in the spatial attention mechanism, which restricts the attention scope to a local region and may not capture disease patterns of varying sizes effectively. Future work could explore adaptive attention mechanisms, such as multi-scale or dynamic receptive fields, to enhance the model's ability to detect features across different spatial contexts.

References

[1] Xinting Yang, Song Zhang, Jintao Liu, Qinfeng Gao, Shuanglin Dong, and Chao Zhou. Deep learning for smart fish farming: applications, opportunities and challenges. *Reviews in Aquaculture*, 13(1):66–90, 2021.

[2] Md Shoaib Ahmed and Samiha Maisha Jeba. Salmonscan: A novel image dataset for machine learning and deep learning analysis in fish disease detection in aquaculture. *Data in Brief*, 54:110388, 2024.

[3] Ming Sun, Xiaofen Yang, and Yinggang Xie. Deep learning in aquaculture: A review. *J. Comput*, 31(1):294–319, 2020.

[4] A Sanjay Kumaar, A Vishnu Vignesh, and K Deepak. Fishnet freshwater fish disease detection using deep learning techniques. In *2024 2nd International Conference on Advancement in Computation & Computer Technologies (InCACCT)*, pages 368–373. IEEE, 2024.

[5] Abdullah Al Maruf, Sinhad Hossain Fahim, Rumaisha Bashar, Rownuk Ara Rumy, Shaharior Islam Chowdhury, and Zeyar Aung. Classification of freshwater fish diseases in bangladesh using a novel ensemble deep learning model: Enhancing accuracy and interpretability. *IEEE Access*, 2024.

[6] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[8] Ananda S Paymode and Vandana B Malode. Transfer learning for multi-crop leaf disease image classification using convolutional neural network vgg. *Artificial Intelligence in Agriculture*, 6:23–33, 2022.

[9] Mohammad Shahjahan Majib, Md Mahbubur Rahman, TM Shahriar Sazzad, Nafiz Imtiaz Khan, and Samrat Kumar Dey. Vgg-scnet: A vgg net-based deep learning framework for brain tumor detection on mri images. *IEEE Access*, 9:116942–116952, 2021.

[10] Zhenbo Li, Fei Li, Ling Zhu, and Jun Yue. Vegetable recognition and classification based on improved vgg deep learning network model. *International Journal of Computational Intelligence Systems*, 13(1):559–564, 2020.

[11] A Victor Ikechukwu, S Murali, R Deepu, and RC Shivamurthy. Resnet-50 vs vgg-19 vs training from scratch: A comparative analysis of the segmentation and classification of pneumonia from chest x-ray images. *Global Transitions Proceedings*, 2(2):375–381, 2021.

[12] Devvi Sarwinda, Radifa Hilya Paradisa, Alhadi Bustamam, and Pinkie Anggia. Deep learning in image classification using residual network (resnet) variants for detection of colorectal cancer. *Procedia Computer Science*, 179:423–431, 2021.

[13] Jiazhi Liang. Image classification based on resnet. In *Journal of Physics: Conference Series*, volume 1634, page 012110. IOP Publishing, 2020.

[14] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR, 2019.

[15] Gonçalo Marques, Deevyankar Agarwal, and Isabel De la Torre Díez. Automated medical diagnosis of covid-19 through efficientnet convolutional neural network. *Applied soft computing*, 96:106691, 2020.

[16] Gonçalo Marques, Antonio Ferreras, and Isabel de la Torre-Diez. An ensemble-based approach for automated medical diagnosis of malaria using efficientnet. *Multimedia tools and applications*, 81(19):28061–28078, 2022.

[17] Md Alamin Talukder, Md Abu Layek, Mohsin Kazi, Md Ashraf Uddin, and Sunil Aryal. Empowering covid-19 detection: Optimizing performance through fine-tuned efficientnet deep learning architecture. *Computers in Biology and Medicine*, 168:107789, 2024.

[18] İsmail Kunduracıoğlu and İshak Paçal. Deep learning-based disease detection in sugarcane leaves: Evaluating efficientnet models. *Journal of Operations Intelligence*, 2(1):321–235, 2024.

[19] Jiangchuan Liu, Mantao Wang, Lie Bao, and Xiaofan Li. Efficientnet based recognition of maize diseases by leaf image classification. In *Journal of Physics: Conference Series*, volume 1693, page 012148. IOP Publishing, 2020.

[20] Mamta Gehlot and Geeta Chhabra Gandhi. "effinet-ts": A deep interpretable architecture using efficientnet

for plant disease detection and visualization. *Journal of Plant Diseases and Protection*, 130(2):413–430, 2023.

[21] Ümit Atila, Murat Uçar, Kemal Akyol, and Emine Uçar. Plant leaf disease classification using efficientnet deep learning model. *Ecological Informatics*, 61:101182, 2021.

[22] Jameer Gulab Kotwal, Ramgopal Kashyap, and Pathan Mohd Shafi. Artificial driving based efficientnet for automatic plant leaf disease classification. *Multimedia Tools and Applications*, 83(13):38209–38240, 2024.

[23] Xuqiang Yin, Dihua Wu, Yuying Shang, Bo Jiang, and Huaibo Song. Using an efficientnet-lstm for the recognition of single cow's motion behaviours in a complicated environment. *Computers and Electronics in Agriculture*, 177:105707, 2020.

[24] Galib Muhammad Shahriar Himel, Md Masudul Islam, and Mijanur Rahaman. Utilizing efficientnet for sheep breed identification in low-resolution images. *Systems and soft computing*, 6:200093, 2024.

[25] Ling Yang, Huihui Yu, Yuelan Cheng, Siyuan Mei, Yanqing Duan, Daoliang Li, and Yingyi Chen. A dual attention network based on efficientnet-b2 for short-term fish school feeding behavior analysis in aquaculture. *Computers and Electronics in Agriculture*, 187:106316, 2021.

[26] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.

[27] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)*, pages 3–19, 2018.

[28] Sandhya Rani Biswal, AVS Pavan Kumar, and Sudheer Babu Punuri. Enhanced fish health monitoring in aquaculture with attention-based deep learning technique. In *2024 International Conference on Integrated Intelligence and Communication Systems (ICIICS)*, pages 1–7. IEEE, 2024.

[29] Naren Sengodan. Cbam-efficientnetv2 for histopathology image classification using transfer learning and dual attention mechanisms. *arXiv preprint arXiv:2410.22392*, 2024.

[30] Subir Biswas, Debendra Muduli, Md Ariful Islam, Anuradha Shantanu Kanade, Abu Taha Zamani, Shantanu Pandurang Kanade, and Nikhat Parveen. Empirical evaluation of deep learning techniques for fish disease detection in aquaculture systems: A transfer learning and fusion-based approach. *IEEE Access*, 2024.

Hoang-Tu Vo holds a Bachelor of Information Systems from Can Tho University, Vietnam, in 2011. In 2014, He graduated with a master's degree in Information Systems from Can Tho University in Vietnam. In 2025, he is a PhD student in Information Systems at Can Tho University, Vietnam. Currently, He is working as a lecturer in Department of Information Technology, FPT University, Can Tho Campus, Vietnam. His research interests include Machine learning, Deep learning, Image processing, and Computer vision.
Email: tuvh6@fe.edu.vn

Huu-Hoa Nguyen received his Engineering Degree in Computer Science from Can-Tho University, Vietnam. He earned his MSc in Information Systems from HAN University, the Netherlands, and his PhD in Informatics from Lyon University, France. Dr. Nguyen is currently a senior lecturer at the College of Information and Communication Technology, Can Tho University, Vietnam. His research interests span a wide range of computer science topics, including artificial intelligence, machine learning, data mining, knowledge management systems, computer networks, and cybersecurity. He has led multiple international research projects, receiving funding from prestigious sources such as the Horizon-2020 European Commission and the Newton Fund
Email: nhhoa@ctu.edu.vn

Dinh Vinh Nguyen (Member, IEEE) received the B.Sc. degree (magna cum laude) in computer science from Nong Lam University, Ho Chi Minh City, Vietnam, in 2007, and the M.S. and Ph.D. degrees in electrical and computer engineering from Sungkyunkwan University, Suwon, South Korea, in 2012 and 2015, respectively. From 2015 to 2019, he was a Research Professor with the College of Information and Computer Engineering, Sungkyunkwan University. From 2019 to 2024, he was a Lecturer and the Head of Artificial Intelligence with the School of Computing and Information Technology, Eastern International University, Vietnam. Since 2024, he has been a Lecturer and a Researcher with the Department of Information Technology, FPT University, Can Tho Campus, Vietnam. His research interests include deep learning, computer vision, image processing, and graphics processing unit computing.
Email: VinhND29@fe.edu.vn